Content based Object Observation for Image Retrieval

Asif Khan
School of Computer Science and Engineering,
UESTC, Chengdu, 611731, China

Jian-Ping Li
School of Computer Science and Engineering,
UESTC, Chengdu, 611731, China

Riaz Ahmed Shaikh
School of Computer Science and Engineering,
UESTC, Chengdu, 611731, China

ABSTRACT
Vision is the most advance of our senses, so it is not surprising that image play the single most important role in human perception. A logical place of overlap of image's processing and their analysis is recognition of object. To extract attributes from image up to include the reorganization of individual object, analysis of text, extracting characters, describing, recognition that and then use in Doman knowledge with analysis of complexity for making sense become vision. This all steps required integration, for that towards vision, from image processing to analysis till goal a new model "SHODHNI". Than model proceed for improvement of pictorial information, storage image data, transmission and representation and then attributes as sense with observation based on task limit and that follow real environment learning. To make that learning possible this model support with offline database also provide and follow full online cloud robo as big data scenario.

Keywords
Artificial Intelligence; Cloud Computing; Big Data; Content Based Image Retrieval; Image Analysis; Vision;

1. INTRODUCTION
Environmental perception and object recognition is an important part of the image processing; it can be widely used in robot visual perception, video surveillance, exception handling, intelligent early warning and mass rapid retrieval and efficient image storage, camera and other fields. Human can perceive the complex scenes easily and respond to get the location and type of the target object correctly, but currently this is a challenging problem for cloud robot visual understanding.

Vision phenomena with natural strategy is a ideal approach and easy task for human being to understand, to see and feel, but that unstructured random scene perception and understanding is a challenging problem in visual image and processing research area. Nowadays image science is a important aspect in education, health, military etc, means image processing under a hyper boom of discovery for this all but to store image, to make it visualize, reduce noise, make enhancement, fix structure divide in pattern all of this with image processing need, only cause of how to make possibility that technological become capable to see and understand.

Due to lack of this ideal platform, judgment of complex scene, obtained the location and try to target object accurately is a complex task for machine or robots in comparison of human being, eyes also have best capability to capture than filter it and after purification that is sense by neurons. In this proposal, proposed model propose a two phase aspect that is provide an intelligent platform for integration of features and preprocess with a second phase of purifier intelligent filter, due to this research work that is inspired by biological vision model possible to purify, segmentation, identification than observation of object become simple and efficient.

Like biological model vision approach always need a platform where easily, properly, fast and efficiently all process can be happened at all features, output of every preprocess on features easily become towards input than all circular loop after every per bit cycle can be generate a proper domain knowledge of identification and observation by which system can easily seen and understand.

Image retrieval systems may needs appropriate object recognition, texture recognition, context retrieval techniques to retrieve information from images.

2. BACKGROUND
To perceive unstructured random capricious natural scene perception and understanding from complex environment is a challenging issue in visual imaging and processing research area. Humans can judge the complex scenes quickly and respond to obtain the location and type of the target object accurately, which is currently the most advanced technological vision systems cannot match. Low level processing of luminous information by the retinal cells. Marr suggested that the coding of this luminous information was based on the zero-crossings of an operator that is now called a wavelet transform. The challenge was to construct a robot endowed with a perception of its environment that enabled it to perform specific tasks. It turned out that the first attempts to construct a robot capable of understanding its surroundings were completely unsuccessful.

Scene understanding is a challenging and difficult problem related to vision applications, which is a key part of the complex natural scenes with unstructured and randomness. Perception of complex natural scenes is difficult to balance and positioning algorithm efficiency and recognition results. Scene associated with each other in different combinations among countryside, adaptive design of the system is difficult to improve understanding of the effect of natural scenes. Therefore, based on visual attention mechanism of natural scene image processing faces enormous challenges.

Environmental perception and object recognition is an important part of the image processing; it can be widely used in robot visual perception, video surveillance, exception handling, intelligent early warning and mass rapid retrieval and efficient image storage, camera and other fields.

Studies have shown that [2], the factors of affecting visual attention from two aspects, i.e., top-down prior knowledge and input signal makes the sensor stimulus from bottom to up. Among them, the top-down prior knowledge and applications of highly correlated, which is very tough for modeling analysis. Therefore, there are a lot of stimulation only sensor modeling of Bottom-up model of visual attention. The bottom –up visual attention model can be divided into two categories

[3]: one is to use the eye tracker eye to gaze position of the image, using statistical methods to get the zone of eye gaze longer time, and as a significant area of interest of mankind. Another type is characterized by multi-channel input image multi-scale analysis, statistically significant distribution of interest on the extent of each pixel in the image and based on the distribution of the extraction are of interest.

Koch and Ullman (1985) [4] firstly proposed the visual attention model based on significant distribution map. Since then, there have been many visual attention model based on significant distribution in [3-7]. Neither of these models needs to deal with complex domain knowledge, does not require the input image to the human eye gaze and gaze time statistics of the number of tedious experiments, which are convenient for application.

These models will be applied to target detection [1,8,9], video compression and coding [10,11], image analysis [12,13] and scene understanding [14] and other fields, which can use the limited memory computing resources to process the input video image or the region of most interest of human vision.

Serreetal in 2005 and 2007 applied HMax model to object recognition [24-25], the improved model constructed high-level simulation of biological visual features. When object recognition, continuous along the hierarchy template matching and merging operation, simulated visual cortex in scale, translation and rotation invariant. There are many researchers in this field have made outstanding contributions [26-30].

The key purpose of CBIR is to locate all images in a particular database that contains convinced visual features stated by the user. The CBIR based image retrieval system, nowadays receiving significant attention which can be implemented more effectively with neural network. Image retrieval systems based on technology for visual image content has been the focus of so many years. Content-based Image Retrieval searches images from large databases by considering the real contents of the digital image (contents refers to shape, texture, color) instead of the metadata e.g. keywords, tags, and/or descriptions related with the image.

3. METHODOLOGY

The major drawback of Image indexing is that while the image is the smallest unit in the image database, it does not lend itself directly to content-based representation. Any given scene is a complex collection of parts or objects. The location and physical qualities of each object, as well as the interaction with others, define the content of the scene. Object-based techniques try to identify objects and relationships among these objects
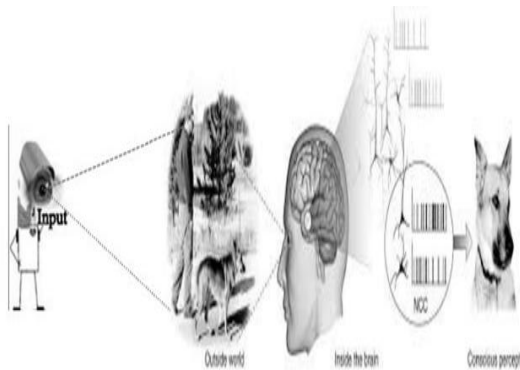

Fig. (1): Real World Approach

In a still image object segmentation and identification is normally a difficult task. In a image sequence, an object moves as a whole. Therefore, we can group pixels that move together into an object. Object segmentation is quite accurate by using the above idea. Object-based image indexing and retrieval can be performed easily when image is compressed using the object-based coding standard
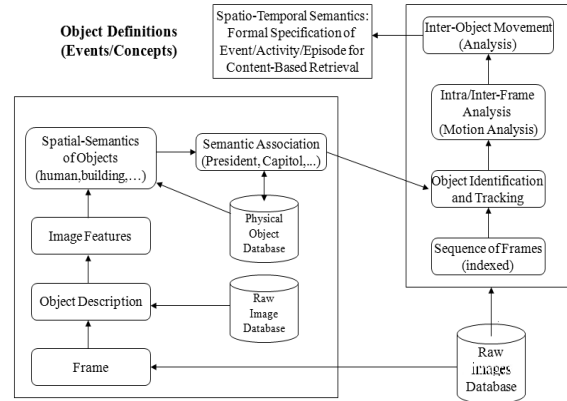

Fig. (2): Approach's Flow

Given an image and object category, to segment the object, Random Fields for segmentation, Generative approach models joint , Markov random field (MRF)

$$p(h\,|\,I,\theta) \propto p(I,h\,|\,\theta) = p(I\,|\,h,\theta)\,p(h\,|\,\theta)$$

create an intelligent relationship between processing of image towards make capable to machine can see. This models primitive layer and comparative layer integrates all image processing features and preprocess by which analysis of image become easier. Nowadays different work environment maintain a huge databases means complex database of image also, so by this model, try to analysis of image on behalf of their preprocess and features . Models goal layer capable to perform an intelligence of identification, observation with their artificial approach, cause of this possibility second phase of this proposed model reduce complexity of vision for machine and robots.
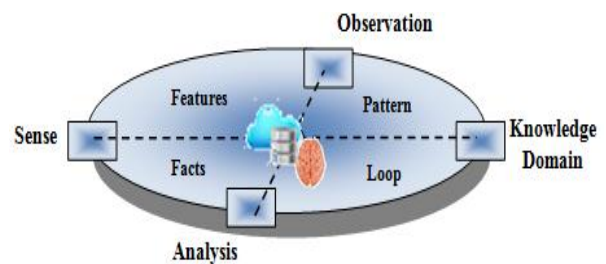

Fig. (3): Domain to Observe

The reorganization of the particular object can be realized by using the comparison with the standard image template stored in database. After simulate the visual cortex structure to construct the object recognition is a viable approach, visual system divide the different image features (e.g., shape, motion, color, spatial position, etc.) into different layer in parallel, each of the hierarchical structure of a serial access. In all parallel processing paths, the three most important steps required from processing a image comparative a image and then observing like sense as object recognition towards complete a vision process like as generative Markov Random Field

$$p(h, I \mid \theta) = p(I \mid h, \theta)\, p(h \mid \theta)$$

Here I (pixels) and h (labels) follow towards {foreground, background}, Pair wise Potential (MRF) ψij(hi, hj|θij) Prior has no dependency on I

$$= \frac{1}{Z(\theta)}\left[\prod_i \phi_i(I \mid h_i, \theta_i) \prod_{ij} \psi_{ij}\left(h_i, h_i \mid \theta_{ij}\right)\right]$$

Dependency on I allows introduction of pair wise terms that make use of image. For example, neighboring labels should be similar only if pixel colors are similar Contrast term

$$p(h \mid I, \theta) = \frac{1}{Z(I, \theta)}\left[\prod_i \phi_i(h_i, I \mid \theta_i) \prod_{ij} \psi_{ij}\left(h_i, h_i, I \mid \theta_{ij}\right)\right]$$

Image retrieval systems may needs appropriate object recognition, texture recognition, context retrieval techniques to retrieve information from images. However, it is a challenge to retrieve proper information from distorted or blurred images. To minimize the complexity of information retrieval from distorted images resembling of images are used in literature. Therefore, to restore the true image, images are often treated as realizations of a random process.
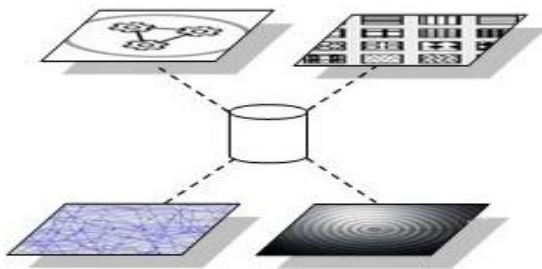

Fig. (4): Third Layer Features Factor

Ω is a shape prior on the labels from a Layered Pictorial Structure (LPS) model Segmentation by, match LPS model to image (get number of samples, each with a different pose marginalize over the samples using a single graph cut

$$p(h \mid \Omega, I, \theta) \propto \left[\prod_i \phi_i^1(I \mid h_i, \theta_i)\, \phi_i^2(h_i \mid \Omega) \prod_{ij} \psi_{ij}^1(h_i, h_j \mid \theta_{ij}) \cdot \psi_{ij}^2(I \mid h_i, h_j, \theta_{ij})\right]$$

For vision processing, image analysis is the important stage; it is the process of extracting significant information from digital images. Image analysis consists of some tasks such as edge detection, shapes detection, objects counting, or measuring object properties. General image analysis algorithms comprise detection of edge, shape detectors, segmentation based on color, and image thresholding.
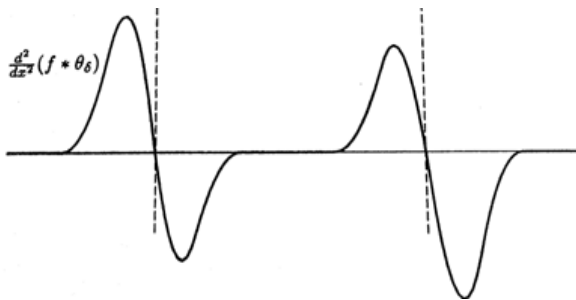

Fig. (5): Observation's Situation

Vision is the most advanced of our sense, so it is not surprising that image play the single most important role in human perception. However unlike humans, who are limited to the visual aspect machine also have no clear cut boundaries in the continuous form, from image processing to computer vision. However according to image processing paradigm, in this proposal image's features and their preprocess functions approach will be integrate

4. CONCLUSION
The scene understanding and recognition by consolidating new intellectual visual features into the scene expression that can be very crucial and provide cognitive intelligence to cloud robot, Inspired by Artificial Neural Network intelligence due to its dynamic nature, we make use of the attributes of SHODHANI which is to be akin to robot visual perception, and apply the cloud computing to inspect a new approach in complex environment natural scene perception and understanding for virtual phenomena. Through the study of Neural Network, the perception ability of the natural scene image from complex environment for cloud robot is enhanced with the integration of cognitive visual features and the scene expression.

5. ACKNOWLEDGEMENT

6. REFERENCES
[1] Najmenik, J., Geisler, W., S., "Optimal Eye Movement Strategies in Visual Search", Nature 434(3), pp. 387-391, 2005.

[2] Siagian, C., Itti, L., "Rapid Biological-Inspired Scene Classification Using Features Shared with Visual Attention", IEEE Trans. on Pattern Analysis and Machine Intelligence, 29(2), pp. 300-312, 2007.

[3] Burt, P., J., Adelson, E., H., "The Laplacian Pyramid As a Compact Image Code" IEEE Trans. on Communication , 31(4), pp. 532-540, 1983.

[4] Afraz, S., R., Kiani, R., Esteky, H., "Micro simulation of Infer temporal Cortex Influences Face Categorization, Nature 442, pp. 692-695, 2006.

[5] Afraz, D., A., Boundar, I., V., Giese, M., A., "Norm Based Face Encoding By Single Neurons in The Mokey Infer temporal Cortex", Nature 442, pp. 572-575, 2006.

[6] Neocognitron, K., F., "A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition unaffected by shift in position Biological Cybemetics. 36(4), pp. 193-202, 1980.

[7] Rolls, E.T., Milward, T., "A Model of Invariant Object Recognition in the Visual System: Learning..

[8] K. Barnard, P. Duygulu, D. Forsyth, N. D. Freitas, D. M. Blei, and M. I. Jordan (2003), "Matching Words and Pictures", the Journal of Machine Learning Research, vol. 3, pp. 1107–1135, 2003.

[9] D. Djordjevic and E. Izquierdo (2007), "An Object- and User-Driven System for Semantic-Based Image

Annotation and Retrieval", Circuits and Systems for Video Technology, IEEE Transactions on, vol. 17, no. 3, pp. 313–323, 2007.

[10] Y. Zhang, C. Xu, YongRui, J. WANG, H. Lu (2007), "Semantic Event Extraction from Basketball Games using Multi-Modal Analysis", Multimedia and Expo, 2007 IEEE International Conference on, 2-5 July 2007.

[11] Najmenik, J., Geisler, W., S., "Optimal Eye Movement Strategies in Visual Search", Nature 434(3), pp. 387-391, 2005.

[12] Itti, L., Koch, C., Niebur, E., "Model of Saliency-Based Visual Attention for Rapid Scene Analysis", IEEE Trans.

on Pattern Analysis and Machine Learning, 20(11), pp. 1254-1259, 1998.

[13] Walther, D., Koch, C., "Modeling Attention to Salient Proto-Objects", Neural Networks, 19(9), pp. 1395-1407, 2006.

[14] Itti, L., Koch, C., "Computational Model of Visual Attention", Nature Reviews Neuroscience, 2(3), pp. 194-203, 2001.

[15] Chen, Z., Han, J., Ngan, K., N., "Dynamic Bit Allocation for Multiple Video Object Coding", IEEE Trans. on Multimedia 8(6), pp. 1117-1124, 2006.